Optimal relay location and power allocation for low SNR broadcast relay channels

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1109/INFCOM.2011.5935119">http://dx.doi.org/10.1109/INFCOM.2011.5935119</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Institute of Electrical and Electronics Engineers (IEEE)</td>
</tr>
<tr>
<td>Version</td>
<td>Author’s final manuscript</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Dec 30 07:35:47 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/73675">http://hdl.handle.net/1721.1/73675</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution-Noncommercial-Share Alike 3.0</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by-nc-sa/3.0/">http://creativecommons.org/licenses/by-nc-sa/3.0/</a></td>
</tr>
</tbody>
</table>
Abstract—We consider the broadcast relay channel (BRC), where a single source transmits to multiple destinations with the help of a relay, in the limit of a large bandwidth. We address the problem of optimal relay positioning and power allocations at source and relay, to maximize the multicast rate from source to all destinations. To solve such a network planning problem, we develop a three-faceted approach based on an underlying information theoretic model, computational geometric aspects, and network optimization tools. Firstly, assuming superposition coding and frequency division between the source and the relay, the information theoretic framework yields a hypergraph model of the wideband BRC, which captures the dependency of achievable rate-tuples on the network topology. As the relay position varies, so does the set of hyperarcs constituting the hypergraph, rendering the combinatorial nature of optimization problem. We show that the convex hull C of all nodes in the 2-D plane can be divided into disjoint regions corresponding to distinct hyperarcs sets. These sets are obtained by superimposing all k-th order Voronoi tessellation of C. We propose an easy and efficient algorithm to compute all hyperarc sets, and prove they are polynomially bounded. Then, we circumvent the combinatorial nature of the problem by introducing continuous switch functions, that allows adapting the network hypergraph in a continuous manner. Using this switched hypergraph approach, we model the original problem as a continuous yet non-convex network optimization program. Ultimately, availing on the techniques of geometric programming and p-norm surrogate approximation, we derive a good convex approximation. We provide a detailed characterization of the problem for collinearly located destinations, and then give a generalization for arbitrarily located destinations. Finally, we show strong gains for the optimal relay positioning compared to seemingly interesting positions.

Index Terms—Low SNR, computational geometry, network optimization.

I. INTRODUCTION

Next-generation wireless standards, such as 3GPP Long Term Evolution-Advanced (LTE-A) standard [1], propose relays as a mean to extend cellular coverage or to increase data rates. More specifically, LTE-A defines relays of Type I as coverage-extension relays which allow a base station (BTS) to reach uncovered users in a cell, and relays of Type II as relays which allow to increase the communication rate of a user already covered through a direct link to the BTS [1], [2]. In terms of cellular deployment, a natural and practical question arises as to where the relay node should be deployed.

In this paper, with the downlink of a cellular system with relays in mind, we address the aforementioned question for the broadcast relay channel (BRC), which consists of a single source broadcasting to multiple destinations with the help of a relay. In this paper, we focus on the wideband regime of wireless relay networks, also denominated low signal-to-noise ratio (SNR) regime because power is shared among a large number of degrees of freedom, making the average SNR per degree of freedom low. We would like to point out that addressing the low-SNR regime is relevant in next generation cellular systems. Indeed, considering LTE, large bandwidths—up to 20 MHz—can be supported by all terminals (3). Due to power constraints in the low SNR regime, relays appear as a meaningful and natural way to increase rate and reliability.

Previous results on wireless systems in the low-SNR regime include the capacity of point-to-point additive white Gaussian noise (AWGN) channel [4], and multipath fading channel [5]–[11], both equal to the received SNR: $C_{\text{Fading}} = C_{\text{AWGN}} = \frac{P}{N_0} = \lim_{W \to \infty} W \log (1 + \frac{P}{W N_0})$; the capacity of the multiple input multiple output (MIMO) channel [12], [13]; the capacity region of the AWGN broadcast channel (BC) [14]–[16], and AWGN multiple access channel (MAC) [17]; and bounds on the capacity of the non-coherent multipath fading relay channel [18]. From these works, a conclusion can be drawn on wireless systems in the low-SNR regime; the major impairment in the low-SNR regime is neither multipath fading nor interference, but noise, which is in contrast with the high-SNR regime. Formulating the argument more concretely, in the presence of multipath fading in the low-SNR, the same rates as the AWGN system with the same received SNR can be achieved using non-coherent peaky signals whereas spread-spectrum signals perform poorly. Moreover, the low-SNR regime is not interference-limited: in particular, all sources in the low-SNR MAC can achieve their interference-free point-to-point capacity to the destination. Based on this observation, the authors proposed in a recent work [19] an equivalent hypergraph model for the low-SNR AWGN MAC and BC. Then they used these models to build an achievable hypergraph model for a more complex wireless network with fixed sources, relays, and destinations, and showed that optimizing power for maximizing multiple session rates boils down to a

ITMANET - 6915101: This material is based upon work under subcontract: 18870740-37362-C, issued by Stanford University and supported by the Defense Advanced Research Projects Agency (DARPA).
straightforward linear program.

In this paper, we take a step forward by simultaneously optimizing the relay location and the power allocation, to maximize the multicast rate from a source to a set of destinations \( T \). Using concepts from information theory, computational geometry and network optimization, we develop a comprehensive and efficient way to solve this problem, which can be broadly divided into three parts:

1) **BRC hypergraph model:** We propose a hypergraph model for the low-SNR BRC, which depends on the topology of the network, essentially the placement of nodes on a 2-D plane. Given the source and destinations positions, the hyperarcs in the BRC hypergraph model requires to get the ordering of nodes in increasing distances from the source and relay, for all relay positions (as the relay is not initially given). This problem can be modeled as an ordered \( k \)-nearest neighbor problem, for which we propose a solution based on superimposing the Voronoi tessellations of all \( k - 1 \) orders, where \( k \) spans the destination set.

2) **Continuous hypergraph variations:** For fixed source and destination positions, the relay position varies, the network hypergraph changes accordingly rendering the problem combinatorial. Consequently, traditional network optimization techniques cannot be applied directly, as they assume a fixed given hypergraph. To circumvent this hard combinatorial nature, we introduce continuous switch functions which allow to change the network hypergraph in a continuous manner as the relay position changes. Ultimately, this allows us to cast the problem as a continuous optimization problem.

3) **Convex approximation:** The resulting continuous network optimization problem is non-convex. However, using geometric programming (GP) and \( p \)-norm approximation techniques, we provide a good convex approximation of the original problem to which standard convex optimization techniques can be applied [20]. It should be noted that the problem is NP-Hard in its original form mainly due to combinatorial nature and continuous non-convex constraints.

Hereafter, the paper is as follows. In Sections II and III, we build the system model and formulate the general problem, respectively. In Section IV we solve the problem for collinearly located destination nodes, and introduce algorithms to compute distinct hypergraphs for various relay positions. Section V extends to the general problem case for an arbitrary topology, finally leading to the conclusion in Section VI.

II. Low SNR system model

Notations: \( \mathbb{N} \) and \( \mathbb{N}^*_m \) denote the sets of non-negative integers, and real numbers, respectively. Let \( m \in \mathbb{N} \) and \( \mathbb{N}^*_m \triangleq \{1, \ldots, m\} \) and let \( S \) be a set, the indicator function of \( S \) is defined by \( 1_S(x) = 1 \) if \( x \in S \), \( 1_S(x) = 0 \) if \( x \notin S \).

In this section, we first recall the equivalent hypergraph models of the wideband BC and MAC, then we use them to build an achievable hypergraph model of the BRC, and finally we formulate the optimization problem. In hypergraph models, a hyperarc \( (u, v_1, v_2, \ldots, v_k) \) of size \( k \) connects a transmitter \( u \) to an ordered (increasing order of distance from transmitter \( u \)) set of \( k \) receivers \( \{v_1, v_2, \ldots, v_k\} \), all of which can decode a message sent over the hyperarc equally reliably. Here, \( u \notin \{v_1, v_2, \ldots, v_k\} \). Two hyperarcs are disjoint if either they have different sources, or different ordered receiver sets or both, e.g. \( (u, v_1, v_2) \) is disjoint from \( (u, v_2, v_1) \) and \( (u', v_1, v_2) \). Messages sent over any pair of disjoint hyperarcs are independent. A hyperarc is said to be activated if its capacity is non-zero.

A. Wideband BC and MAC model

Earlier results on multiple user channels show that BC and MAC are not impaired by interference in the low SNR regime.

1) **Equivalent hypergraph of the wideband AWGN BC.** [12]–[21]: Superposition coding is known to achieve the capacity region of the AWGN BC. In the wideband limit, the rates achieved by superposition coding boils down to the time-sharing rates, rendering time-sharing as optimal.

Consider the BC channel with source \( s \), two destinations \( T = \{d_1, d_2\} \) in Figure 1(a), \( D_{sv} > 0 \) is the distance of node \( v \) from \( u \), and let \( \alpha \in [0, 1] \) be the power-sharing factor at source. Then both destinations can receive the common rate \( \alpha \min\{h_1^2, h_2^2\} \frac{P_{\text{no}}}{N_0} \), and the most reliable destination can also receive a bonus private rate \( (1 - \alpha) \max\{h_1^2, h_2^2\} \frac{P_{\text{no}}}{N_0} \), where \( h \) is the path loss factor, \( N_0 \) the channel noise and \( P \) is the total source power. This motivates the equivalent hypergraph model of the wideband AWGN BC in Figure 1(b). The wideband hypergraph model contains three hyperarcs: the common hyperarc from the source to \( d_1 \) and \( d_2 \) with capacity equal to the private rate \( R_{\text{com}} \), a private edge from \( s \) to \( d_1 \) with capacity equal to the private rate \( R_{\text{com}} \) if \( d_1 \) is more reliable than \( d_2 \), i.e. \( h_1^2 > h_2^2 \), and to 0 otherwise, and finally a private edge \( s \) to \( d_2 \) with a capacity equal to the bonus rate \( R_{\text{com}} \) if \( d_2 \) is more reliable than \( d_1 \). Note that the two private hyperarcs (associated with rates \( R_{\text{com}} \) and \( R_{\text{com}} \)) cannot exist simultaneously (as either \( D_{s1} < D_{s2} \) or \( D_{s2} < D_{s1} \)): thanks to the indicator functions in the capacity expressions, only one of the private hyperarcs can be activated for a given topology. In the general case of a wideband AWGN BC with \( n \) destinations:

- For an arbitrary unknown topology, the full hypergraph model contains \( 2^n - 1 \) hyperarcs, from the source to every possible subset of destinations.
- For a given known topology, only a subset of these hyperarcs are activated simultaneously. Indeed, a given topology yields a given ordering of the destination set in increasing order of reliability. Consequently, only \( n \) hyperarcs are simultaneously activated for a given topology: one private arc of size 1 to the most reliable destination i.e. \((s, d_1)\), and one common hyperarc of size \( k \) to the \( k \) most reliable destinations for all \( k \in \{2, \ldots, n\} \) i.e. \((s, d_1, \ldots, d_k)\).

2) **Equivalent hypergraph of the wideband AWGN MAC:** Consider two sources \( s_1 \) and \( s_2 \) and a single destination \( d \) in Figure 1(c). In the wideband regime, the large number of degrees of freedom renders negligible interference, and
allows all sources to achieve their point-to-point capacity to the destination, as with frequency division multiple access (FDMA). Thus, the respective capacities of $s_1$ and $s_2$ are $C_1 = h_1^2 \alpha_{s_1 r} \frac{P_s}{N_0}$ and $C_2 = h_2^2 \frac{P_s}{N_0}$. As shown in Figure 1(e) the equivalent hypergraph model contains only two edges, one from $s_1$ to $d$ with capacity $C_1$ and one from $s_2$ to $d$ with capacity $C_2$. In the general wideband MAC with $n$ sources, the hypergraph model consists of $n$ hyperarcs of size 1 with non-zero capacity, from each source to the destination.

### B. Wideband BRC model

Consider the broadcast relay channel in Figure 1(e), where a source $s$ transmits to a set of $n$ destinations $T = \{d_i\}_{i \in \mathcal{N}}$ with the help of a relay $r$. We assume that all nodes are equipped with a single antenna. The source and the relay have given respective average power constraints $P_s$ and $P_r$, and they transmit in two different frequency bands, $W_s$ and $W_r$, respectively, so as to respect the half-duplex constraints at the relay. During each time slot, $s$ transmits a new codeword which is received by the relay $r$ relay and $T$ (all destinations); $r$ processes the signal received from $s$ in the previous time slot and retransmits it to $T$; the destinations use the signals they received directly from $s$ and through $r$ to decode a new codeword.

The wireless link between two nodes $u \in \{s, r\}$ and $v \in \{r, d_1, \ldots, d_n\}$ is modeled by an AWGN channel. In other words, when node $u$ transmits a signal $x_u(t)$, node $v$ receives a signal $y_{uv}(t) = h_{uv} x_u(t) + z_v(t)$ where $h_{uv} = \frac{1}{d_{uv}^\alpha}$ is an attenuation coefficient modeling pathloss, and $z_v(t)$ is a white Gaussian noise process with power spectral density $N_0$. Note that although we consider AWGN channels, the low-SNR analysis could be extended to multipath fading channels; indeed, it was shown in [18] that in the wideband multipath fading relay channel, the same rates can be achieved as in the wideband AWGN relay channel with the same average received SNR on each link.

The AWGN BRC consists of two BC components in series for $s$ and $r$, the BC from $s$ to $\{r, d_1, \ldots, d_n\}$ in red in Figure 1(e) of the BRC and the BC from $r$ to $\{d_1, \ldots, d_n\}$ in blue—and of $n$ parallel MAC components, such as the MAC from $\{s, r\}$ to $d_1$ represented by the sum of the red and blue lines resulting into a green line.

As in [19], we make the assumption that the source $s$ and the relay $r$ are constrained to transmit using the scheme that would be optimal for their respective wideband BC-components: $s$ transmits using superposition coding in the source band $W_s$; $r$ decodes the messages it received from $s$, and then retransmits using superposition coding in the relay band $W_r$; each destination $d_i$ decodes by using the interference-free signals it received from $s$ and $r$.

Under these constraints on the communication scheme, the resulting hypergraph model of the BRC [19] is simply the concatenation of the equivalent hypergraphs of the BC-components and the MAC components. We will denote this hypergraph as $G(\mathcal{N}, \mathcal{H})$, where $\mathcal{N} \supset \{d_1, \ldots, d_n\}$ and $\mathcal{H}$ is the set of hyperarcs. The set $\mathcal{H}$ can be partitioned into two disjoint sets: one is $\mathcal{H}_s$ of source hyperarcs emanating from $s$, and the set $\mathcal{H}_r$ of relay hyperarcs emanating from $r$, where $\mathcal{H}_s \cup \mathcal{H}_r = \mathcal{H}$. Figure 1(d) illustrates the hypergraph in the case of a given topology with two destinations. In this figure, we assume that $r$ is the closest node to $s$, followed by $d_1$ and then $d_2$, and we show only the activated hyperarcs.

It should be pointed out that this BRC hypergraph model is only an achievable model, and not an equivalent model. Indeed, in the case of a single destination, the BRC boils down to the relay channel, and it was shown in [13] that with a different coding scheme, it is possible to achieve a higher rate in the wideband relay channel than any rate obtained by the scheme in [19]. Thus the BRC hypergraph model proposed in this paper provides only an achievable rate region, but not the full rate region of the BRC. However the relaying scheme in [19], and the associated hypergraph model, have the benefit to easily extend to large complex network.

### III. General problem structure

Given a topology of the set of nodes $\mathcal{N} \setminus \{r\}$, and the aforementioned achievable hypergraph model of the wideband BRC, we recall the following question: What is the optimal relay position and power allocations at $s$ and $r$, that maximize the multicast rate $R_m$ from $s$ to the destination set $T$? Here, the multicast rate is the rate experienced by the least reliable destination in the set $T$, and is given by its min-cut. To solve the problem in full generality, we propose a two-stage approach, as follows:

- **Pre-processing:** The pre-processing stage computes all distinct $\mathcal{H}_s$ and $\mathcal{H}_r$, respectively, for all positions of the relay inside the region being considered on the 2-D plane, given by hyperarcs $\mathcal{H}_s$ and $\mathcal{H}_r$, respectively. Since,
only a subset of these hyperarcs are active when the relay is in a certain region, we associate each hyperarc \((u, V) \in \mathcal{H}\), along with a continuous switch function \(f_{uv}\). The switch function \(f_{uv}\) activates the hyperarc \((u, V)\) by taking the value 1 when it should exist and deactivates the hyperarc by taking value 0 when it should not exist. In this section and section IV, we devise efficient algorithms to compute all the distinct hyperarcs. Once the hyperarc set \(\mathcal{H}\) is constructed, we can then compute all the possible paths from the source to each destination. The total number of paths from \(s\) to a destination \(d_i \in T\) will be denoted \(K_i\), and the rates on these paths will be written \(\{r^1_{di}, ..., r^K_{di}\}\).

- Optimization: The second stage involves solving a network flow optimization problem. After obtaining \(\mathcal{H}\), the multicast rate maximization problem for \(G(N, \mathcal{H})\) can be formulated as:

\[
\text{Program (A): maximize } (R_m) \\
\text{subject to: } R_{rn} \leq r_i, \forall i \in N_n, \quad (1) \\
r_i \leq \sum_{k=1}^{\infty} r^{di}_{ki}, \forall i \in N_n, \quad (2) \\
\max_{(i, k) : k_i \in [1, K_i]} r^{di}_{ki} \leq y_{uv}, \forall i \in N_n, R_{uv} \in H, \quad (3) \\
y_{uv} \leq c_{uv} f_{uv}, \forall (u, V) \in H, \quad (4) \\
\text{where, } c_{uv} \in C_{uv}, \forall (u, V) \in H. \quad (5)
\]

Here \([1]\) implies \(R_m\) is the minimum among the total rates experienced at all destinations, \([2]\) says that the rate for destination \(d_i\) is the sum of rates on all the paths from \(s\) to \(d_i\). \([3]\) captures the network coding constraint, and the switch function \(f_{uv}\) in \([4]\) activates and deactivates hyperarc \((u, V)\) as the optimization algorithm goes from one relay position to the other to maximize \(R_m\). \([5]\) implies that the capacity of hyperarc \((u, V) \in \mathcal{H}\) is determined by implicit constraints of power and distance.

Note that, it is because of continuous switch functions that we have a continuous optimization problem, else we would need combinatorial constraints to capture the right hyperarcs that need to be activated for each relay position.

In the sequel of this section, we describe the two stages in detail. At first, we show that the optimal relay position lies in the convex hull \(C\) of nodes \(N \setminus r = \{s, T\}\).

A. Convex hull lemma

We prove hereunder an intuitive lemma which helps in formulating the problem in a geometric sense.

**Lemma 1:** Given \(G(N, \mathcal{H})\), the relay location that maximizes the multicast rate \(R_m\) from \(s\) to \(T\) lies inside the convex hull of the nodes \(N \setminus r\).

**Proof:** We prove this lemma by building an intuitive argument. To start with, consider the four node system of \(\{s, r, d_1, d_2\}\) in Figure 2. The only node that is allowed to take its desired location is the relay \(r\). The convex hull \(C\) of the node set \(N \setminus r = \{s, d_1, d_2\}\) is shown as the shaded area in Figure 2. Consider the arbitrary position outside \(C\) in the above scene e.g. \(r\) is placed outside \(C\). The hyperarc sets for this position of \(r\) are given by \(H_s = \{(s, r, 1), (s, d_1), (s, r, d_1, d_2)\}\) and \(H_r = \{(r, d_1), (r, d_1, d_2)\}\), where \(\mathcal{H}\), assuming that the ordered sets of nodes in increasing distances from \(s\) and \(r\) are given by \((s, d_1, d_2)\) and \((r, d_1, d_2)\), respectively.

Consider the rates on the path \(\{(s, r, (r, d_1)\}\) from \(s\) to \(d_1\), which are given by:

\[
R_{sr} \leq \frac{P_{sr}}{D_{sr} N_0}, \quad R_{r1} \leq \frac{P_{r1}}{D_{r1} N_0}, \quad (6)
\]

where, \(P_{sr}\) and \(P_{r1}\) are source and relay powers. It is clear that by moving \(r\) towards the boundary of \(C\), i.e. line segment joining \(s\) and \(d_1\), total rate on this path \((\min(R_{sr}, R_{r1}))\) could be increased. The triangle’s inequality corroborates this fact,

\[
D_{sr} + D_{rd1} \geq D_{sd1}, \quad (7)
\]

\[
(R_{sr} + R_{r1})(D_{sr} + D_{rd1}) \geq D_{sd1}, \quad (8)
\]

\(7\) is the triangles inequality for the triangle \(\Delta_{srd1}\) and this implies \(8\), which states that the rate from \(s\) to \(d_1\) on the path \(\{(s, r), (r, d_1)\}\) could be increased by simply bringing the \(r\) to towards the line segment joining \(s\) and \(d_1\). It is straightforward to see that this also increases the rate for all other receivers in system (consider triangles \(\Delta_{srd1}\)).

This reasoning can easily be generalized to any arbitrary position of the relay outside the convex hull \(C\), and to any arbitrary number of destinations \(|T| > 2\). Thus, we conclude that for any given instance of BRC, the relay location maximizing the multicast rate lies inside or on the border of the convex hull of \(\{s + T\}\) nodes, but never outside. Hence, proved.

**Lemma 2** implies that we only need to consider relay locations in the convex hull \(C\) of nodes \(N \setminus r\), to maximize the multicast rate. There are efficient algorithms for constructing a convex hull of \(n + 1\) points, c.f. [22] and references within.

B. Pre-processing algorithms

The pre-processing stage consists of three sub-stages. For all relay positions in \(C\), the first sub-stage computes the source hyperarc set \(\mathcal{H}_s\), the second sub-stage computes the relay hyperarc set \(\mathcal{H}_r\), and the last sub-stage computes all the source-destination paths \(\forall d_i \in T\). Now, we develop algorithms...
to compute $\mathcal{H}_s$ and $\mathcal{H}_r$ and give upper-bounds on the number of distinct hyperarcs in $\mathcal{H}_s$ and $\mathcal{H}_r$.

1) Source hyperarcs ($\mathcal{H}_s$): We first prove a lemma on $|\mathcal{H}_s|$. 

**Lemma 2:** The number of distinct source hyperarcs inside the convex hull $C$ is upper bounded by $3n - 1$, where $n = |T|$. 

**Proof:** Consider a BRC with $s$ and an ordered set of three destinations $T = \{d_1, d_2, d_3\}$ ($|T| = n = 3$). Let $c_i$ be the circle centered at $s$ passing through $d_i$. An example is illustrated in Figure 3(a), the three circles partition the 2-D plane into rings and discs that are given by disc $C_1$ and two concentric rings $R_{21}$ and $R_{32}$.

For the positions of $r$ inside these areas there are distinct sets of source hyperarcs. Computing the source hyperarcs when $r$ is inside these regions and starting with the disc $c_1$, we simply get a set of 4 hyperarcs $H_s = \{(s, r), (s, rd_1), (s, rd_1d_2), (s, rd_1d_2d_3)\}$. Each time $r$ crosses the border of circle $c_i$ and enters the ring $R_{2i-1}$, there are two hyperarcs that change and hence the new hyperarcs must be added, i.e., $\{(s, d_1), (s, d_1r)\}$ for $R_{21}$ and $\{(s, d_1d_2), (s, d_1d_2r)\}$ for $R_{32}$, (refer Figure 3(b)). This is due to the fact that, when $r$ enters a new region, the ordered set of nodes in increasing distances from $s$ of the new region is different from the previous region in only two places, e.g. for the disc $C_1$ and ring $R_{21}$ the ordered sets are given by $(s, r, d_1, d_2, d_3)$ and $(s, d_1, r, d_2, d_3)$, respectively.

Thus, the maximum number of distinct source hyperarcs that can exist for all relay positions in $C$ is given by $(n + 1) + 2(n - 1) = 3n - 1$. Hence, proved.

At this point we would like to highlight a couple subtleties: 
(a) If some $d_i \in T$ are equidistant from $s$, then their respective circles coincide, hence reducing the number of disjoint rings, and the number of distinct hyperarcs becomes less than $3n - 1$. Thus, Lemma 2 is an upper bound. 
(b) The simple algorithm outlined in Lemma 2 builds all the possible hyperarcs efficiently in the sense that only distinct hyperarcs are added to $\mathcal{H}_s$ along with their respective switch functions, thus avoiding any redundancy.

**Switch Function:** The activation/deactivation of a hyperarc can be performed by the switch function. For instance, the switch function associated with hyperarc $(s, rd_1)$ (Figure 4) is $f_{sr1} = (1 + \gamma e^{(-\gamma z_{sr1})})^{-1}$, where $z_{sr1} = D_{s1} - D_{sr}$ and $\gamma >> 1$. When $r$ is in $C_1$, $z_{sr1}$ is positive, and $f_{sr1} \approx 1$, thus hyperarc $(s, rd_1)$ is active. Similarly, when $r$ is outside $C_1$, $z_{sr1}$ is negative, and $f_{sr1} \approx 0$, thus hyperarc $(s, rd_1)$ is deactivated. Notice, that for the hyperarcs inside a ring (e.g. $R_{vi-1}$), the switch function will be a product of two functions, each for the regions of concentric circles that make this ring.

2) Relay hyperarcs ($\mathcal{H}_r$): The relay hyperarcs are determined by the ordering of node set $T$ with increasing distance from $r$. Thus, we need to partition the 2-D plane into disjoint regions where the ordering of $T$ with respect to $r$ stays the same. This is equivalent to computing the order-$k$ Voronoi tessellations of $C$ for all $k \in \mathbb{N}_{n-1}$, and then superimposing them to get the ordered set $T$ of nearest neighbors in $C$. (22) It is known that the superimposition of Voronoi tessellations results in convex disjoint areas (polygons in our case). Once the ordered set $T$ for each disjoint region is computed, the $n$ activated relay hyperarcs for each region are obtained the same way as in Section II-A for the case of a BC with a given known topology. Thus, the algorithm outlined in Lemma 1 could be used to generate $\mathcal{H}_r$. Figure 4 illustrates the superimposed disjoint regions of ordered destinations with respect to $r$ for $n = 4$. The simplest way to compute these regions (ref. 24) is to draw the perpendicular bisector of every destination pair $(d_i, d_j)$ in $T^2$. This method has the complexity of order $O(n(n-k))$, where $k \in \mathbb{N}_{n-1}$. Hence, we obtain the partitions of $C$ for distinct ordering of the set $T$ with respect to $r$ from which we can generate $\mathcal{H}_r$.

The pre-processing in almost all network planning problems is computationally heavy as there is plenty of time up-front compared to real-time applications. In our case it fits better as all the computations are of polynomial order.

In the next sections, we formulate the problem of optimal relay positioning as a non-convex network flow optimization problem and provide a good convex approximation. For simplicity and clarity in understanding, we divide the problem into two cases. The first case is for collinearly located destination nodes and the second case is for arbitrarily located destinations.
IV. COLLINEAR CASE

In this section we develop the method to solve the simpler version of the problem where the destination nodes are collinearly located. It helps understanding the main concepts and underlying algorithms, and ultimately leads to the solution for the arbitrary case.

A. Pre-processing

1) Convex hull $C$: For the collinearly located destination set $T$, the set $T$ could be ordered in the increasing order of abscissa, for instance with the left most node being at the origin. Hereafter, we assume that the left-most (respectively rightmost) node $d_l \in T$ ($d_r \in T$) is situated at the origin (respectively right most at horizontal axis), and the rest of the destination nodes are on the positive x-axis, and finally that the source is in the positive quadrant. Note that, $s$ could be the leftmost node compared to any $d_l \in T$, in this case $s$ could be assumed to be on the positive vertical axis and the set $T$ would accordingly be placed on positive horizontal axis with $d_l$ not being at origin. Since all $d_l \in T$ are collinearly located, the convex hull will always be the triangle $\triangle sd_l d_r$. (ref. Figure 5(a)). Thus, $C$ is always given by only three inequalities in this case.

2) Source hyperarcs: As explained in Section III-B1 the source hyperarcs are functions of the source-destination distances. Consequently, the algorithm outlined in Lemma 2 could be used to compute $\mathcal{H}_s$.

3) Relay hyperarcs: Also, as shown in Section III-B2 we compute the perpendicular bisectors of every destination pair $(d_i, d_j) \in T^2$, to compute the superimposed convex disjoint regions of ordered destination sets w.r.t. $r$ (ref. Figure 5(b)). For the collinearly located nodes, the computation of the set $\mathcal{H}_r$ is greatly simplified due to parallelism of all bisectors. The following Lemma is just an easy and straightforward formalization.

Lemma 3: For $|T| = n$ collinear destinations, the total number of distinct relay hyperarcs in $C$ is upper bounded by $n^2 = n + 2(\beta - 1)$, where $\beta = \binom{n}{2} + 1$ is the number of bisected regions.

Proof: Given a set $T$ of $n$ collinear destinations, the maximum number of bisectors are given by $\binom{n}{2}$. Then the total number of bisected regions is given by $\beta = \binom{n}{2} + 1$, as shown in Figure 5(b). Since, crossing each bisector only changes two nodes in the ordered destination set, using the algorithm outlined in Lemma 2 (ref. Figure 3), we can compute all distinct relay hyperarcs to $T$, $|\mathcal{H}_r| = n + 2(\beta - 1) = n^2$. Hence, proved.

With only little more formalization of Lemma 3 we can device easy algorithms for computing the set $\mathcal{H}_r$. Note, that the switch function for each relay hyperarc can be computed in a similar manner as for the source hyperarcs. At the same time, a particular switch function could constitute two sub-switch functions each for two perpendicular bisectors.

4) Source-destination paths: After successfully computing all distinct source ($\mathcal{H}_s$) and relay ($\mathcal{H}_r$) hyperarcs for a given system, we now need to compute all the paths from $s$ to all

Fig. 5. Pre-processing. The triangle $\triangle sd_2 d_3$ shows $C$, with circles and perpendicular bisectors dividing $C$ is closed and disjoint sets. (a): Shows ordered 4-tuple set for each region carved by the circles $c_1, c_2$ and $c_3$ as $C_1, C_2, C_3$ and $C_4$, respectively. (b): Shows the ordered 3-tuple sets of destination nodes with respect to $r$. (c): Shows the previous 2 figures superimposed showing the disjoint convex regions with the ordered sets of closest nodes with respect to $s$ and $r$, respectively. The 2-tuple $(C, r)$ represents the ordered sets for each region, respectively. Here, $r_1, r_2, r_3$ and $r_4$ represents the four regions in (b).
destinations $d_i \in T$, in order to successfully cast our problem as a network optimization program. We prefer a path based formulation as opposed to a more basic and standard link based formulation because the path based formulation is far well suited for convex approximations of originally non-convex network optimization programs, in our framework.

There are many efficient ways (polynomial time algorithms) to compute the paths from the set $H_s$ and $H_r$. For simplicity, we prefer here to take all combinations of the hyperarcs in $H_s$ and $H_r$. Not all these paths will be active for a certain relay position, but the switch functions will take care of activating/deactivating the paths. In this way, we get an upper bound $(3n - 1)(n^2) = |H_s| \times |H_r|$ on the paths from $s$ to each $d_i \in T$, out of which only a certain number of paths will have non-zero min-cut, i.e. activated hyperarcs. We define, $\Omega = [1, (3n - 1)(n^2)]$ as the set of all paths from $s$ to every $d_i \in T$. This makes the problem size bigger, but saves cost of activated path computation. Note that the total number of paths are polynomially bounded in our model.

\subsection*{B. Optimization}

In this section we formalize the problem of optimal relay location maximizing the multicast rate from $s$ to $T$. The optimization constraints can be grouped into two categories: the posynomial constraints that can be easily rewritten using exponential transformation as convex constraints (Geometric Programming); and the non-posynomial constraints that can only be approximated as convex constraints. Since, all the constraints are coupled through variables, almost all the variables and hence constraints will go through the exponential transformation. Now, we classify and discuss the troubling constraints in our formulation.

1) Hyperarc rate constraints (Posynomials): We show an example of hyperarc rate constraints. Consider the hyperarc $(s, d_1, r_2)$ in the scenario in Figure 5(a) which is active when the relay is inside the ring $R_{21}$. The non-convex rate inequality can be expressed as:

$$y_{s,1,2} \leq \frac{P_{s,1} f_{s,1,2}}{D_{s,1} N_0} f_{s,1,2}$$

where $f_{s,1,2} = f_{s,1}^2 f_{s,2} f_{s,1,2}^2$, $f_{s,1,2}^2 \leq \left(1 + \gamma e^{(\gamma z_{s,1,2})^2}\right)^{-1}$, $f_{s,1,2}^2 \leq \left(1 + \gamma e^{(\gamma z_{s,1,2})^2}\right)^{-1}$, $D_{sr} - D_{s1} = z_{s,1,2}$ and $\alpha_{s,1,2} = D_{s2} - D_{sr}$. Notice, how $f_{s,1}^2$ and $f_{s,2}^2$ and their respective $z$ variables are different. When, $f_{s,1}^2$ and $f_{s,2}^2$ take the value as 1, the hyperarc will have a non-zero min-cut. Rewriting them together,

$$\frac{R_{s,1} D_{s,1} N_0}{P_{s,1} f_{s,1,2}} \leq 1, \quad \frac{R_{s,2} D_{s,2} N_0}{P_{s,2} f_{s,1}^2 f_{s,2}} \leq 1,$$

where $f_{s,1,2} = f_{s,1}^2 f_{s,2} f_{s,1,2}^2$, $f_{s,1,2}^2 \leq \left(1 + \gamma e^{(\gamma z_{s,1,2})^2}\right)^{-1}$, $f_{s,1,2}^2 \leq \left(1 + \gamma e^{(\gamma z_{s,1,2})^2}\right)^{-1}$, $D_{sr} - D_{s1} = z_{s,1,2}$ and $\alpha_{s,1,2} = D_{s2} - D_{sr}$.

2) Distance function constraints (Non-posynomials): Variables $D_{uv}$ in rate inequalities represents distance functions, given by e.g.,

$$\sqrt{(x_r - x_s)^2 + (y_r - y_s)^2} = D_{sr},$$

where $(x_s, y_s)$ are fixed coordinates of $s$ and $(x_r, y_r)$ are the variable coordinates of $r$. The negative coefficients in (14) prohibits the use of GP techniques. There are techniques to get around this problem [25], [26], but the extra pre-processing cost incurred is very high in addition to the introduction of many new variables and combinatorial constraints.

We prefer to handle the issue in a simpler manner by approximation. Let, the only variable transformed using GP in (14) be $D_{sr}$. Then, we can rewrite

$$u_{sr}^2 + v_{sr}^2 \leq e^{2D_{sr}}^r, \quad x_r - x_s \leq u_{sr}, \quad y_r - y_s \leq v_{sr}.$$  

The first inequality in (15) is non-convex. Using the $p$-norm surrogate approximation (27) for (19), we get

$$\left(\frac{u_{sr}^2 + v_{sr}^2}{e^{2D_{sr}^r}}\right)^p \leq 1,$$

where $p \in [1, +\infty)$. Over a compact set of variables and in the limit of $p \to \infty$, (16) becomes convex. In our case, for values of $p = 4$ or 5, we get good approximation. Note, that only the first inequality in (15) needs to be approximated, and since the variables $(u_{sr}, v_{sr}, x_r, y_r)$ don’t undergo GP transformation, the rest of the inequalities in (15) remain linear.

All other constraints in the program are posynomials, as we will shortly see, so they can easily be transformed into convex constraints. It should be noted, that it is only because of the use of switch functions that the program becomes continuous. In addition, carefully designing the switch function results in posynomial hyperarc rate constraints.

3) Network Optimization problem formulation: Since there are $\Omega = (3n - 1)(n^2)$ number of paths for each destination $d_i \in T$ only a subset of them will actually be active (i.e. with min-cut > 0), $r_{d_i}^q$ as the rate on path $q$ to destination $d_i$, where $q \in \Omega$. Recall, $\mathcal{H} = H_s \cup H_r$ and the total rate to a destination $d_i$ be defined as $r_i = \sum_{q \in \Omega} r_{d_i}^q$. Also, let $v_m \in V$ be the farthest node from $i$ for hyperarc $(u, V) \in \mathcal{H}$. Then, the optimization program is,

Program (B): maximize $(R_m)$ subject to:

$$R_m \leq r_i, \forall i \in \mathbb{N}_n,$$

$$r_i \leq \sum_{q \in \Omega} r_{d_i}^q, \forall d_i \in T,$$

$$\max \frac{R_i^q}{((u,V) \in \mathcal{H})} \leq \frac{P_{uV}}{D_{uV} N_0} f_{uV}, \forall (u, V) \in \mathcal{H},$$

$$f_{uV} \leq \frac{f_{uV}^1 f_{uV}^2}{\gamma e^{(\gamma z_{uV})^2}}, \forall (u, V) \in \mathcal{H},$$

$$f_{uV} \leq \frac{1 + \gamma e^{(\gamma z_{uV})^2}}{\gamma e^{(\gamma z_{uV})^2}}, \forall (u, V) \in \mathcal{H},$$

$$l \in [1, 2], \forall (u, V) \in \mathcal{H},$$

$$u_{uV}^2 + v_{uV}^2 \leq D_{uV}^2, \forall (u, V) \in \mathcal{H},$$

$$\sum_{((u,V) \in \mathcal{H})} P_{uV} \leq P_s, \sum_{((u,V) \in \mathcal{H})} P_{uV} \leq P_r.$$
where \( x_r - x_u \leq u_{uv}, \forall (u, V) \in \mathcal{H}, \) \( 25 \)
\[ y_r - y_u \leq v_{uv}, \forall (u, V) \in \mathcal{H}, \]
\( x_r \geq 0, y_r \geq 0, y_r \leq \lambda x_r, y_r + \lambda x_r \leq \eta. \) \( 27 \)

In the above program \((x_r, y_r)\) are variable relay coordinates and (27) captures constraints that make \( C \). Program (B) is a non-convex program expressed in posynomial and polynomial inequalities. Applying GP transformation to the following variables \( \{ r_{uv}^d, P_{uv}, D_{uv}, f_{uv} \} \), \( p \)-norm approximation to constraints \( 23 \) and leaving the rest of the variables unchanged, we get the following convex approximation,

\[
\begin{align*}
\text{Program (C): maximize } & \min_{d_i \in T} \left( \sum_{q \in \Omega} r_{id_i}^q \right) \\
\text{subject to: } & N_0 e^{(r_{uv}^c + \alpha D_{uv}^c - f_{uv}^c)} \leq 1, \\
& \forall q \in (u, V), \forall (u, V) \in \mathcal{H}, \\
& e^{(f_{uv}^c - f_{uv}^c)} \leq 1, \forall (u, V) \in \mathcal{H}, \\
& e^{(f_{uv}^c - f_{uv}^c)} + e^{(f_{uv}^c - f_{uv}^c)} \leq 1, \forall (u, V) \in \mathcal{H}, \\
& x_{s1} \leq D_{uv}, \forall (u, V) \in \mathcal{H}, \\
& \left( \frac{u_{uv}^2 + v_{uv}^2}{e^{D_{uv}}} \right)^p \leq 1, \forall (u, V) \in \mathcal{H}, \\
& \sum_{(u, V) \in \mathcal{H}_s} e^{p_{uv}} \leq P_s, \sum_{(u, V) \in \mathcal{H}_r} e^{p_{uv}} \leq P_r,
\end{align*}
\]

where, we have used the GP transformation \( x^c = \log(x) \) (\( x \) is the original variable of program (B)) on certain variables.

Program (C) is a convex approximation of program (B) with no underlying combinatorial hard structure. The approximation is only coming from constraint type (32) using \( p \)-norm surrogation technique that gives a convex approximation to the constraint \( 23 \) in program (B).

Note, that the objective function is modified in (C), instead of having the sum of positive exponential terms, we have replaced it by a sum of linear functions, which is far easier to maximize. The maximizers of the function \( \max_X(x_1 + x_2) \) also maximizes the function \( \max_X(e^{x_1} + e^{x_2}) \), over a compact set \( X \) with certain particular characteristics. This is generally not true, but in our case due to network coding constraints with certain simple tricks it can be proven that it holds true. Due to the lack of space we omit the detailed proof.

As we know, with the increasing value of \( p \), the program approaches a complete convexity with zero duality gap, thus standard convex optimization algorithms could be used to solve problem (C) with increasing accuracy. The optimal values of program (B) could be easily constructed from the optimal values of program (C).

C. Simulations

In this section, we present simulation results for the BRC with \( n = 2 \) destinations. We compare the multicast rate obtained by optimizing the relay location and the source in addition to relay power allocations, with the case where the relay is located at a naive yet seemingly interesting position: the centroid of triangle \( \triangle_{s_1, d_2} \), and only the power allocations are optimized. The simulations are run for an increasing size of the area of \( \triangle_{s_1, d_2} \) and a random network topology for each area is chosen.

Figure 6 shows the maximum multicast rate (blue and red) for optimal and centroid relay positions respectively. The SNR \( \frac{P_0}{N_0} \) is normalized to 1. Note that the actual values of the rates are not as important because of the normalization, for higher power values, the rate would certainly have higher values. For increasing area of triangle \( \triangle_{s_1, d_2} \), the maximum multicast rate tends to drop, which is due to the constrained power and larger distances, but the relative gain goes up. This implies that for farther placed nodes the sensitivity of the relay location is higher and can produce significant gains in rate for the optimal relay location. Its clear from the results in Figure 6 that the centroid is not the optimal location. The rise in the relative gain becomes very strong due to the fact that the low-SNR regime is more sensitive to the location of nodes (hence, distances) that determine the hyperarc rates in the limit of disappearing SNR as opposed to e.g. in high SNR regime, where a displacement of \( \pm \epsilon \) for the location of \( r \) would not effect the rate significantly.

V. ARBITRARY CASE

In this section, we answer the same set of questions but for arbitrarily placed source and destination nodes. Almost all concepts can be carried over to, straightforwardly.

The steps of the pre-processing stage can be summarized as Input: \( \{s, T\} \) set of nodes with their cartesian coordinates.

1) Compute convex hull \( \mathcal{C} \).
2) Compute \( \mathcal{H}_s \), switch functions (using Lemma 2).
3) Compute the disjoint convex regions by superimposing all \( k \)-order Voronoi diagrams of \( \mathcal{C} \).
4) Compute \( \mathcal{H}_r \) (using relay hyperarc algorithm).

Output: \( \mathcal{H} = \mathcal{H}_s \cup \mathcal{H}_r \).

Once we have \( \mathcal{H}_s \) and \( \mathcal{H}_r \), we can compute \( \Omega = [1, |\mathcal{H}_s|] \times [|\mathcal{H}_r|] \). Ultimately, the optimization program could be stated as,

\[
\begin{align*}
\text{Program (D): maximize } & (g_0(x)) \\
\text{subject to: } & g_i(x) \leq 1, \ i \in [1, k], \\
& g_j(x) \leq 1, \ j \in [k + 1, K],
\end{align*}
\]

where, the constraints \( 34 \) are the posynomials constraints that can be transformed to convex convex constraints using GP and constraints \( 35 \) are the non-posynomial constraints that are approximated using \( p \)-norm approximation. The objective
function \( g_0(x) \) represents the multicast rate. In program (D), \( x \) is a vector of variables.

Program (D) is an abstract representation of the actual program. Since, the program (D) is simply program (B) (but for arbitrary placement of destination nodes), the structure of (D) is the same as (B). The main difference is the pre-processing stage for the two cases, in this case which involves computation of \( k \) nearest neighbor nodes and superimposing them to form disjoint regions in \( C \) for distinct \( n \)-nearest neighbors. We would like to note, that this computation, although polynomially bounded, can be heavy. There are many polynomial time algorithms in the literature of computational geometry that solve this problem efficiently, \([23]\).

VI. RESULTS AND CONCLUSION

A comprehensive and efficient solution is developed to model and answer the problem of optimal relay positioning so as to maximize the multicast rate from the source \( s \) to the destination set \( T \) in a low-SNR network. The proposed solution is a non-convex network optimization problem in its basic form that is difficult to solve. Using GP, switch functions and \( p \)-norm surrogate approximation we transform the problem to a convex approximation that can be solved using standard convex optimization algorithms.

To abridge, the important contributions of this work could be summed up in the following words: we present a comprehensive approach to determine the optimal relay position under the pretext of network optimization problem. Network topologies consisting single source, multiple destinations with the only intermediate node as relay are considered in complete generality on a 2-D plane. Using superposition coding and frequency division we construct a wireline like hypergraph. The low-SNR hyperarc model using superposition coding provides an interference free network model that is easily scalable to complex network topologies. Using the tools of computational geometry and network optimization, we presented a network optimization framework based solution that is intuitive and easy to understand. Also, we show that positioning the relay optimally significantly affect the network performance.

In addition, the main causes for complexity in our approach are the pre-processing stage and the non-convexity arising from non-posynomial constraints upon GP transformation. The former reason could be somewhat tolerated, as generally for solving network planning problems heavy pre-processing is required. In contrast, in our case the pre-processing stage consists of polynomial time operations at the cost of only slight sub-optimality in approximation.

The questions our work answers are just a fraction of the interesting questions that it opens up. An interesting direction would be to extend this model to general multimmodity flow optimization problems involving more number of relay nodes. On the other hand, from the computational point of view, an interesting question is how we can efficiently build the exact number of paths essentially bringing down the size of the network optimization program. Finding other techniques to model this problem could be interesting, e.g. utilizing geometric properties of the problem and ways to bring down the computational complexity.

REFERENCES